

Evaluation of face PAD solutions

A bank's journey

Brian Greenough and Max Dermann
Bank of New Zealand



Awarded World's Most Innovative Digital Bank

classification: Public

Bank of New Zealand

Using face biometric matching since 1861

Sensor Mk1 Mod 0 eyeball

Algorithm Human cognition

Struggling with

pose

lighting

artefacts

...



HON. JAMES WILLIAMSON
DIRECTOR BANK OF NEW ZEALAND

Evaluating face PAD solutions

... not quite since 1861

Why is a bank evaluating face Presentation Attack Detection solutions?

BNZ has been moving from face-to-face biometrics to online biometrics.

Online, we can no longer rely on the veracity of any information.

Initial testing showed that static images are trivial to manipulate.

So off we went, into a number of design and test cycles ...

Scenario in focus

The most relevant scenario to BNZ is **online face matching**.

Examples:

Identity attribute verification via a customer's laptop and browser

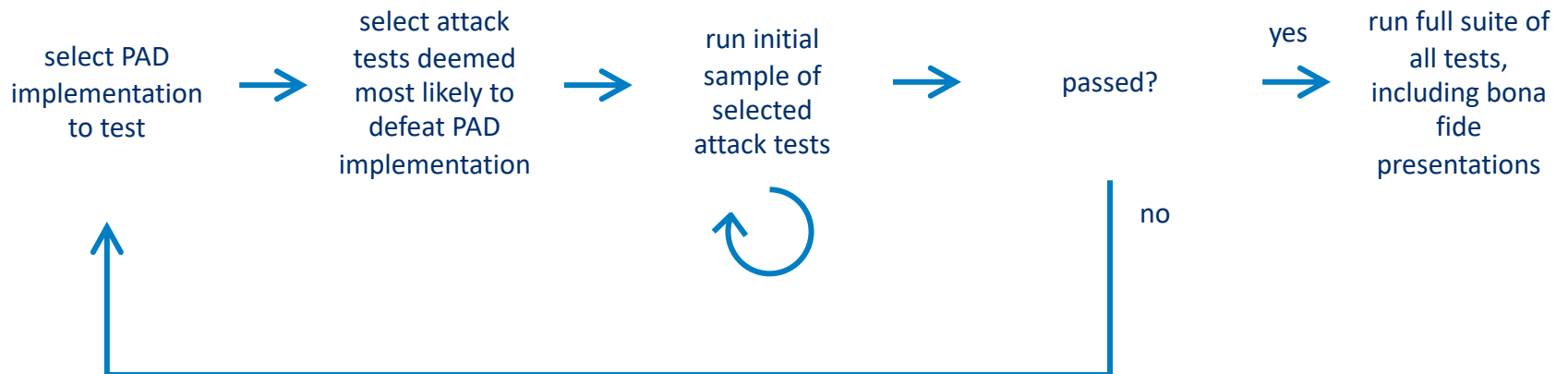
Authentication via a customer's smartphone and BNZ's banking app

Challenge: The images we receive can be

replayed	[(3D-)printouts, displayed on screen, fed into the software]
manipulated	[partial physical masks, digital overlays]
generated	[digital animations]

The BNZ PAD testing methodology

... has been optimised to **economically deliver sufficient certainty** on whether a PAD implementation can reduce impersonation risks to within the BNZ risk appetite.



Selecting the initial tests

... thinking like an attacker

Select those tests that are deemed most likely to overcome the specific PAD implementation.

Guidelines

- G1 Always use the presentation attack technique with the highest level of sophistication for that attack type
- G2 Don't bother testing presentation of static images where the PAD solution is designed to look for movement
- G3 Don't bother testing presentations that do not include the correct responses to any requested challenges

Fine tuning tests

BNZ has identified a large number of variables in the tests that we have developed.

Variables include lighting, reflections, pose and movement (speed, direction).

A systematic approach, that limits variation and iterates through all permutations, requires excessive effort, from a commercial point of view.

We opted to use non-systematic variations (akin to fuzzing or lock picking), which are introduced by the tester.

This relies on the skills and experience of the testers, but has allowed us to efficiently detect fundamental limitations in all assessed implementations.

We call this approach fine tuning.

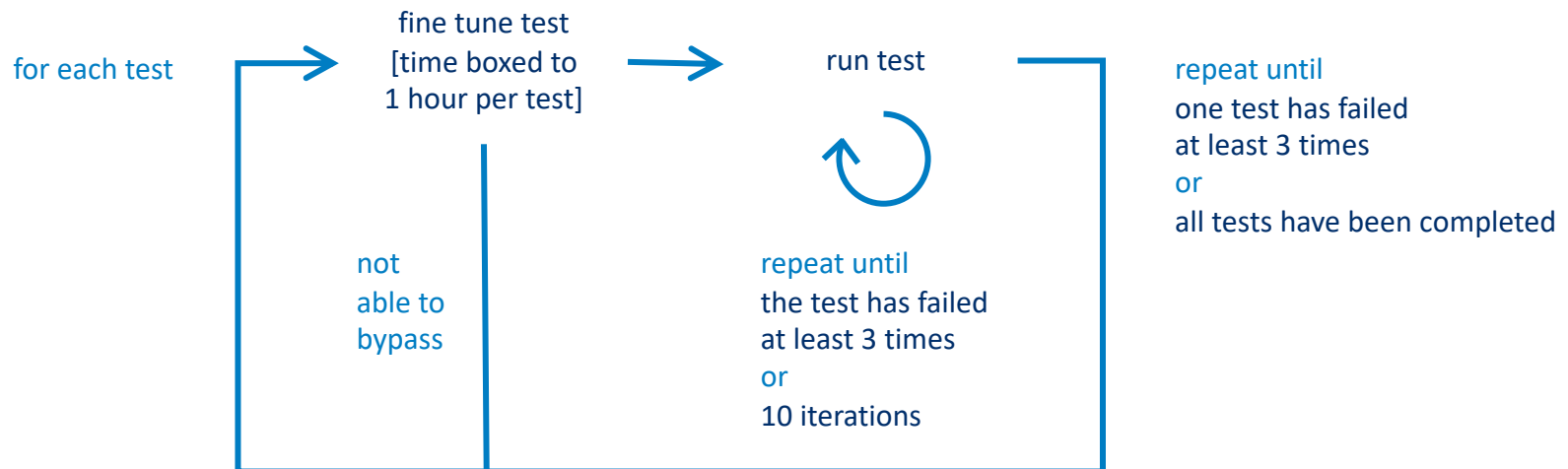
We have limited fine tuning per test to a maximum of one hour [time boxing].

Passed or failed

From a commercial perspective, it is sufficient to understand whether a PAD implementation can be bypassed by at least one attack technique.

Therefore, **testing is stopped after the first failed series of tests.**

If failure can be reproduced in at least 3 out of 10 iterations, a test series has failed.



Tested PAD implementations

... include:

head movement detection (incidental and challenge- response)

facial movement (incidental and challenge- response)

image texture analysis

distance distortion

Note:

BNZ has signed NDAs with most vendors, meaning we can only share anonymised test results. We therefore encourage everyone to validate our results.

Test suite

BNZ has developed a comprehensive face PAD test suite.

The suite contains attack tests as well as bona fide presentation samples.

Tests are continuously refined.

The most effective attack technique has been real-time digital animation, using a COTS animation software (Reallusion CrazyTalk 8, RRP \$149).



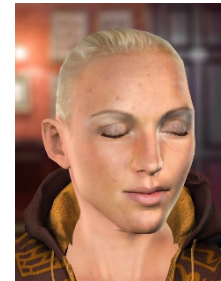
original



animation



original

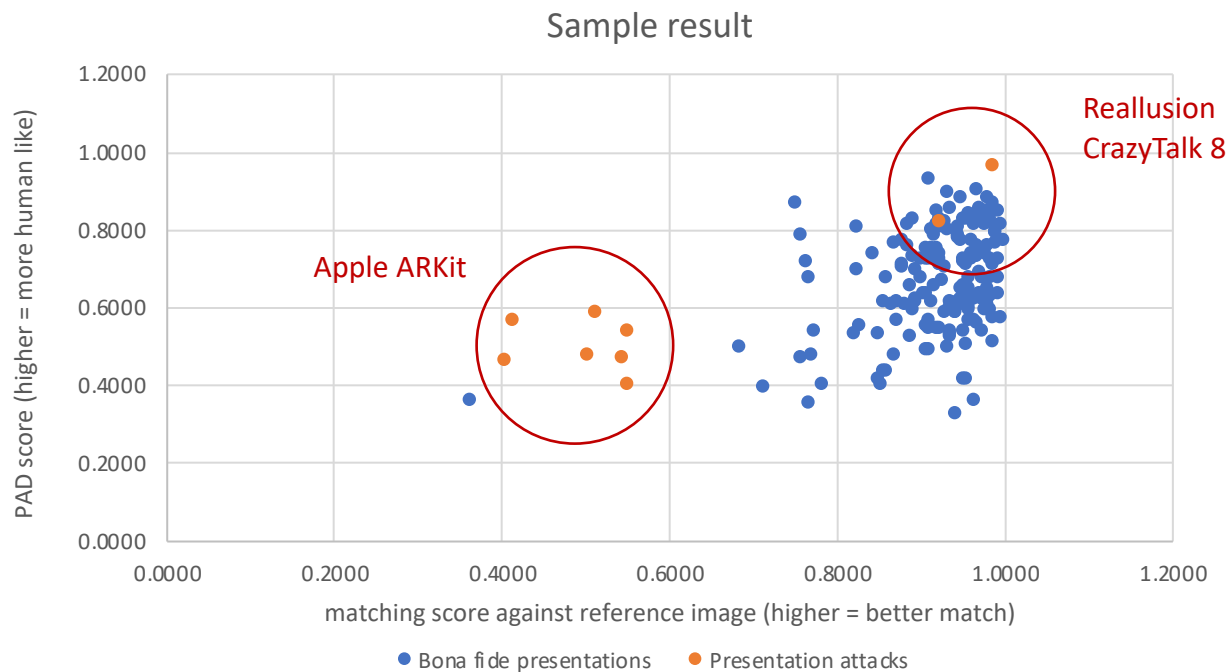


animation

Measurements

BNZ has tested all types of PAD detection implementations that we could identify.

None of the implementations that we have tested managed to pass our tests.



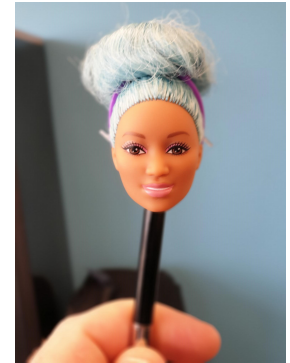
Observations

All tested PAD implementations, apart from one based on distance distortion could be overcome using digital animations.

The tested implementation based on distance distortion could be overcome using video replay.

A tested implementation based on head movement could also be overcome using a small scale 3D sample of a head.

In case of the tested implementations based on image texture analysis, the false rejection rate for images of humans wearing heavy makeup were consistently high.



Interpretation

Digital animation technique has reached a point where it is **hard to impossible to automatically differentiate between real humans and animations when relying on a single, uncontrolled visible light camera.**

We would not be surprised if the total global investment into animation technology by the entertainment industry vastly exceeded that into PAD technologies.

We therefore expect animation technology to reach a point within the next few years where even humans will struggle to identify the differences.

BNZ is currently limiting face biometric applications to controlled environments (e.g. branches, kiosks) and human review (back to Mk 1 Mod 0 eyeball).

While we are not aware of any COTS, low cost animation technology that can imitate other effects like distance distortion and pulse, we assume replicating these effects digitally poses a low technical hurdle.

The limitations of ISO/IEC 30107-3

... or any other security testing standard.

Attack detection and prevention technology is constantly evolving.

Attack technology and techniques are also constantly evolving.

Any security testing standard would either also need to evolve at the same speed or only describe testing techniques on a very high level.

ISO/IEC 30107 has opted for the latter, relying predominately on the skills and imagination of the testers.

Security testers, with their limited time and budgets will always struggle to emulate hundreds if not thousands of real world attackers.